



IFWO

RAW SEQUENCE LISTING

DATE: 09/20/2004

PATENT APPLICATION: US/10/762,769

TIME: 16:31:11

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\09202004\J762769.raw

4 <110> APPLICANT: Melis, Anastasios
 5 Wintz, Hsu-Ching Chen
 7 <120> TITLE OF INVENTION: MODULATION OF SULFATE PERMEASE FOR
 8 PHOTOSYNTHETIC HYDROGEN PRODUCTION
 11 <130> FILE REFERENCE: BERK-016CIP
 13 <140> CURRENT APPLICATION NUMBER: 10/762,769
 14 <141> CURRENT FILING DATE: 2004-01-21
 16 <150> PRIOR APPLICATION NUMBER: 60/354,760
 17 <151> PRIOR FILING DATE: 2002-02-04
 19 <150> PRIOR APPLICATION NUMBER: 60/377,902
 20 <151> PRIOR FILING DATE: 2002-05-02
 22 <150> PRIOR APPLICATION NUMBER: 10/350,298
 23 <151> PRIOR FILING DATE: 2003-01-22
 25 <160> NUMBER OF SEQ ID NOS: 9
 27 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 29 <210> SEQ ID NO: 1
 30 <211> LENGTH: 411
 31 <212> TYPE: PRT
 32 <213> ORGANISM: Chlamydomonas reinhardtii
 34 <400> SEQUENCE: 1
 35 Met Glu Arg Val Cys Ser His Gln Leu Ala Ser Ser Arg Gly Arg Pro
 36 1 5 10 15
 37 Cys Ile Ala Gly Val Gln Arg Ser Pro Ile Arg Leu Gly Thr Ser Ser
 38 20 25 30
 39 Val Ala His Val Gln Val Ser Pro Ala Gly Leu Gly Arg Tyr Gln Arg
 40 35 40 45
 41 Gln Arg Leu Gln Val Val Ala Ser Ala Ala Ala Ala Ala Phe Asp
 42 50 55 60
 43 Pro Pro Gly Gly Val Ser Ala Gly Phe Ser Gln Pro Gln Gln Gln Leu
 44 65 70 75 80
 45 Pro Gln Gln His Pro Arg Gln Pro Gln Ala Val Ala Glu Val Ala Val
 46 85 90 95
 47 Ala Glu Ser Val Ser Ala Pro Ala Ser Ala Ala Pro Ser Asn Asp Gly
 48 100 105 110
 49 Ser Pro Thr Ala Ser Met Asp Gly Gly Pro Ser Ser Gly Leu Ser Ala
 50 115 120 125
 51 Val Pro Ala Ala Ala Thr Ala Thr Asp Leu Phe Ser Ala Ala Ala Arg
 52 130 135 140
 53 Leu Arg Leu Pro Asn Leu Ser Pro Ile Ile Thr Trp Thr Phe Met Leu
 54 145 150 155 160
 55 Ser Tyr Met Ala Phe Met Leu Ile Met Pro Ile Thr Ala Leu Leu Gln
 56 165 170 175
 57 Lys Ala Ser Leu Val Pro Leu Asn Val Phe Ile Ala Arg Ala Thr Glu

(pg.6)

ENTERED

RAW SEQUENCE LISTING

DATE: 09/20/2004

PATENT APPLICATION: US/10/762,769

TIME: 16:31:11

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\09202004\J762769.raw

```

58          180          185          190
59 Pro Val Ala Met His Ala Tyr Tyr Val Thr Phe Ser Cys Ser Leu Ile
60          195          200          205
61 Ala Ala Ala Ile Asn Cys Val Phe Gly Phe Val Leu Ala Trp Val Leu
62          210          215          220
63 Val Arg Tyr Asn Phe Ala Gly Lys Lys Ile Leu Asp Ala Ala Val Asp
64 225          230          235          240
65 Leu Pro Phe Ala Leu Pro Thr Ser Val Ala Gly Leu Thr Leu Ala Thr
66          245          250          255
67 Val Tyr Gly Asp Glu Phe Phe Ile Gly Gln Phe Leu Gln Ala Gln Gly
68          260          265          270
69 Val Gln Val Val Phe Thr Arg Leu Gly Val Val Ile Ala Met Ile Phe
70          275          280          285
71 Val Ser Phe Pro Phe Val Val Arg Thr Met Gln Pro Val Met Gln Glu
72          290          295          300
73 Ile Gln Lys Glu Met Glu Glu Ala Ala Trp Ser Leu Gly Ala Ser Gln
74 305          310          315          320
75 Trp Arg Thr Phe Thr Asp Val Val Leu Pro Pro Leu Leu Pro Ala Leu
76          325          330          335
77 Leu Thr Gly Thr Ala Leu Ala Phe Ser Arg Ala Leu Gly Glu Phe Gly
78          340          345          350
79 Ser Ile Val Ile Val Ser Ser Asn Phe Ala Phe Lys Asp Leu Ile Ala
80          355          360          365
81 Pro Val Leu Ile Phe Gln Cys Leu Glu Gln Tyr Asp Tyr Val Gly Ala
82          370          375          380
83 Thr Val Ile Gly Thr Val Leu Leu Leu Ile Ser Leu Val Met Met Leu
84 385          390          395          400
85 Ala Val Asn Gln Leu Gln Lys Leu Ala Arg Lys
86          405          410

```

89 <210> SEQ ID NO: 2

90 <211> LENGTH: 3873

91 <212> TYPE: DNA

92 <213> ORGANISM: Chlamydomonas reinhardtii

94 <400> SEQUENCE: 2

```

95 gcttagtacc taagcaaaaa taccaaagcc ttatcctgag ttgtcaacaa gaactccagc 60
96 ctgcgacgat gcaaagcctt tcttgagcgg gttgatggac tttgctttgt tatctgtcca 120
97 gtaagccacc agacactacc aagtagagta atccatttgt ataggtacag aatatggagc 180
98 gagtttgcag ccatacagctt gcctcgtcgc gagggaggcc atgcatcgct ggggtgcagc 240
99 ggtcgcccat ccgactaggg acttcaagcg ttgctcatgt gcaggtctct ccggcaggta 300
100 agcaccgcgc tcggcggcgt gtacacatgg gcccgtcagg ccaactgcgt ttgttggtta 360
101 tgcaaccgaa acaggccttg ggagatatca acggcaaaaga ctgcaagtcg tggcgtctgc 420
102 agctgcgcca gcggctttcg accctcctgg aggtgcgtgg cgtgagggct gcacgggtgc 480
103 gggttggcct ggaaaccaag cctcgccacg actacctgca acagcattgc ccgcatctcc 540
104 agccccctcac cctcgagtgc ctcccgaaga cctctatccc ctgcgcatca ttggttcggg 600
105 ggcgcgcgct gcgggccttg ggcgtggct acgctgaccg cacggcacga cttggcacgg 660
106 cctggcgcgg cctgagcggc cccccccctc ctgatggccc cacgctttgc cgcccacgcc 720
107 gctccccgca ggtgtctcgc ccgggttctc gcagccgcaa cagcagctgc cacaacagca 780
108 cccacgccaa ccacaggcgg tggcggaggt agctgtcgcc gagtcagtct cggcgcccgc 840
109 ttctgcggcg ccctccaatg atggctcgcc cacggcctcc atggacggcg gcccagctc 900

```

RAW SEQUENCE LISTING

DATE: 09/20/2004

PATENT APPLICATION: US/10/762,769

TIME: 16:31:11

Input Set : D:\Seqlist.txt

Output Set : N:\CRF4\09202004\J762769.raw

```

110 cggcctcagc gccgtgcccg ccgcgcgcac cgccaccgac ctcttctccg ccgcggcgcg 960
111 cctccgcctg cccaacctct ccccatcat cacttgacc ttcagtctct cctacatggc 1020
112 cttcatgtc atcatgccc ataccgcgt gctgcaaaaa gcctcgctcg tgccgctcaa 1080
113 cgtcttcac gcgcgcgcca ccgagccggt ggcatgac gcctactacg tcaccttctc 1140
114 ctgctcgctg atcgcgccg ccatcaactg cgtgtttggc ttcgtgctgg cctgggtgct 1200
115 ggtgcgctac aatttcgcgg ggaagaagat cctggacgcg gcggtggacc tgccgttcgc 1260
116 gctgccgacc tcggtggcg gcctcacgct tgccacggtg tacggcgacg agttcttcat 1320
117 cggccagttc ctgcaggcgc agggcggtgca ggtgcgtgct tatagcatag tggagtgtgg 1380
118 ttagcagctg ggggtccggc agtagttccc gccctagtga ggtcgaaact ataccagaag 1440
119 aagaggacga acatggggct atccagcaag ctctgtctagg gaaggaggag tttgggagaa 1500
120 cgggtgggtg ggagggagag ggagggcggt ggctgggagg gaagggtgag gcgggagggg 1560
121 gatggttagc cggggcggtg gggacgcaga aggatgacag gcggtgacg ggaagggatg 1620
122 gggaagcgga gctggggaca gtgcgaagag ccgggagaga ggggaagttt gactcaggaa 1680
123 gaggggctag agaggggcat gcggactcct gctgggattt aggtgcgtgc tcattgagga 1740
124 gcccttgaa tcagcggacg gaaacgtggc cgacggggtc tgccgagcac accaggctag 1800
125 ctagacgcgc ggttgggcaa cgagcagagc tgctgtgcgg ctatggatgg aaggcgatgc 1860
126 agcgagcatg tgcagtgaac attggtttga ggacagggga ctccgaggtt gcataggcgg 1920
127 gccgccactg tctctgccgc tagggtgact agctgcctcg aacctggcgg tggccccata 1980
128 ccgcagttg gaggatgtct caccgccttc agcttgccat gtctggggtc tgggtctgga 2040
129 cgcaatcagc gtgtgaggg ccaactctat atggaattat ggataccttc caactaccag 2100
130 cactaggtc gccggaacgc ggtgaagcg gctggcctgc cccctcatcc tctcgttccc 2160
131 ctgtttttgt cccctgtcca ccaggtggt gttcacgcgg ctgggtgtgg tgatcgccat 2220
132 gatcttcgtg tcttccccct tcgtggtgct caccatgcag cccgtcatgc aggtgagagc 2280
133 gccaggagg cggagccatg gcgggttggg gcgggttggg gcgggttggg gcggggcgcg 2340
134 gatggggcg cttggggagt aatgtggggc ggatgggggt gcagcctggc agggatggg 2400
135 agcgagagga tagcggggac aggggacagg gaagggaagg gaagggaagg gatgccctat 2460
136 gcgagcaaag ggggtatggg aaccggcggt tggggctggg agcgacggga gcaggagggg 2520
137 agtgcacgga acgggggcaa ggcggacagg gtgagggagg gtgcaggccg gactgggatg 2580
138 ggtcatgtgt cctggtcggg ggtgtagccg tgggagcgcg gcaggcagcg tgtgttctgg 2640
139 caggtgttt tggcgaaaga taccacggca tggatgggg ccagttgggc agggagaaga 2700
140 cgttgacac gacttcgttg acagatctag ttcattgcac ccgggtcgca ccaagggtgg 2760
141 cggcgagccc ggcccggcac gtccgagtac cccggagccg taacgcccga acccgccctg 2820
142 ttgcgcccc tccctgtctc cctgtctcgc ataccgtgca ccatgccctc tgccgcccc 2880
143 tcaggccctc aggcctcac ctccccctca cctcctccta acgccttccc ctgccttcc 2940
144 ctccccctcc caacgccacc acgtgcaaca ggaaatcaa aaggagatgg aggagggcgg 3000
145 atggtcgctg ggcgcctcg agtggcgcac cttcacagac gtggtgctgc cgccgctgct 3060
146 gcccgcgctg ctgaccggca cggcactggc cttctcgcgc gcgcttggcg agttcggtac 3120
147 cattgtcatc gtgtcctcca actttgcctt caaggacctg atcgcgccc tgctgatctt 3180
148 ccagtgcctg gagcagtag actacgtgg cgccaccgtg atcggcacag tactgctgtt 3240
149 gatttcgctg gtgatgatgt tggcggtgaa ccagctgcag aagctggcgc gcaagtgagg 3300
150 ggctgaggcg tttgaggaga gtgggcgtct gcggaggcgc ttgtggcgca ggggcagggtg 3360
151 gaggaggtt cagggtgagg caggagtggc aggtggtgga ggtgagggg cggggtgttg 3420
152 ggatgggatg ggatgggacc gtgggagggg tgggactttg ggtgggtggg agtgggtgct 3480
153 acgtattagg atatgggagg tggatgacg ttgaagggg ggttgcaat ctggacgggg 3540
154 actcactgtt tactaggcac gcatgtcgca ggagtggata tcgatgggtg tggggatgtc 3600
155 agcacgcttg gcttgagttg ggccatggga cccgggacta ggcttggtg cgagccgagc 3660
156 cagtcaccag ggagacgtac gagcgacac agtgattacg gggattgatt aggcggcgaa 3720
157 ttgacgcaaa tccacggggg ctgtggcttg ggggaggcag ggattgagcg aaggacgcac 3780
158 tgcaagctca ggcagtcgca tgcccgtacc ctgcttctgg tccagtgtgg agacaagact 3840

```

RAW SEQUENCE LISTING

DATE: 09/20/2004

PATENT APPLICATION: US/10/762,769

TIME: 16:31:11

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\09202004\J762769.raw

```

159 ggcaatcgtg gtcctttgca attcatggcg cgc 3873
161 <210> SEQ ID NO: 3
162 <211> LENGTH: 1984
163 <212> TYPE: DNA
164 <213> ORGANISM: Chlamydomonas reinhardtii
166 <400> SEQUENCE: 3
167 gcttagtacc taagcaaaaa taccaaagcc ttatcctgag ttgtcaacaa gaactccagc 60
168 ctgcgacgat gcaaagcctt tcttgagcgg gttgatggac tttgctttgt tatctgtcca 120
169 gtaagccacc agacactacc aagtagagta atccatttgt ataggtacag aatatggagc 180
170 gagtttgacg ccacacgact gcctcgtcgc gagggaggcc atgcatcgct ggggtgcagc 240
171 ggtcgcccat ccgactaggg acttcaagcg ttgctcatgt gcaggctctc ccggcaggcc 300
172 ttgggagata tcaacggcaa agactgcaag tcgtggcgct tcgagctgcy gcagcggtct 360
173 tcgacctcc tggaggtgtc tccgcgggt tctcgcagcc gcaacagcag ctgccacaac 420
174 agcaccacag ccaaccacag gcgggtggcg aggtagctgt cgcagagtcg gtctcggcgc 480
175 ccgcttctgc ggcgcctcc aatgatggct cgcacacggc ctccatggac ggcggcccca 540
176 gctccggcct cagcgccgtg cccgcgcgc ccaccgccac cgacctctc tccgcgcgcg 600
177 cgcgcctccg cctgcccac ctctcccca tcatcacctg gaccttcag ctctcctaca 660
178 tggccttcat gctcatcatg cccatcacgc cgtgctgca aaaagcctcg ctctgcccgc 720
179 tcaacgtctt catcgcgcg gcaccgagc cgggtggcgat gcacgcctac tacgtcacct 780
180 tctcctgctc gctgatcgcg gcgcacctca actgctgtt tggcttcgtg ctggcctggg 840
181 tgctggtgcy ctacaatttc gcggggaaga agatcctgga cgcggcggtg gacctgccgt 900
182 tcgctgctgc gacctcggtg gcgggcctca cgcttgccac ggtgtacggc gacgagttct 960
183 tcatcgcca gttcctgcag gcgcagggcg tgcaggtggt gttcacgcgc ctgggtgtgg 1020
184 tgatcgccat gatcttcgtg tcttccccct tcgtggtgcy caccatgcag cccgtcatgc 1080
185 aggaaatcca aaaggagatg gaggaggcgc catggtcgct gggcgccctc cagtggcgca 1140
186 ccttcacaga cgtggtgctg ccgcgcgtgc tgcccgcgt gctgaccggc acggcactgg 1200
187 ccttctcgcy cgcgcttggc gagttcggat ccattgtcat cgtgtcctcc aactttgcct 1260
188 tcaaggacct gatcgcgccc gtgctgatct tccagtgcct ggagcagtac gactacgtgg 1320
189 gcgcaccgt gatcggcaca gtactgctgt tgatttcgtt ggtgatgatg ttggcggtga 1380
190 accagctgca gaagctggcg cgcaagttag gggctgaggg gtttgaggag agtgggcgtc 1440
191 tgcggaggcg cttgtggcgc aggggcaggt ggaggaggtt gcagggtgag gcaggagtgg 1500
192 caggtggtgg aggggtgcag gcggggtgtt gggatggat gggatggag cgtgggaggg 1560
193 gtgggacttt ggggtgggtg gagtgggtgc tactatttag gatatggag gtggtatgca 1620
194 gttgaagggg ggggtggcaa tctggacggg gactcactgt ttactaggca cgcagtgcgc 1680
195 aggagtggat atcgatgggt gtggggatgt cagcacgctt ggcttgagtt gggccatggg 1740
196 acccgggact aggcttgggt gcgagccgag ccagtcacca gggagacgta cgagcgca 1800
197 cagtgattac ggggattgat taggcggcga attgacgcaa atccacgggg gctgtggctt 1860
198 gggggaggca gggattgagc gaaggacgca ctgcaagctc aggcagtcgc atgcccgtac 1920
199 cctgcttctg gtccagtgtg gagacaagac tggcaatcgt ggtcctttgc aattcatggc 1980
200 gcgc 1984
202 <210> SEQ ID NO: 4
203 <211> LENGTH: 1863
204 <212> TYPE: DNA
205 <213> ORGANISM: Chlamydomonas reinhardtii
207 <400> SEQUENCE: 4
208 cattcaattt gcagcgttcc taaaatggca agcacaacgc tgctccagcc cgcgcttggg 60
209 ctgcctcgc gggtagggcc tcgtccctct ctgtcgcttc ccaaaattcc tcgcgtgtgc 120
210 acgcacacta gtgtccctc tacctcaaag tactgcgact catcatcagt tatagagagc 180
211 acgctagggc ggcaaacatc ggttgccggg agaccatggc ttgcaccccg gcctgcgcct 240

```

RAW SEQUENCE LISTING

DATE: 09/20/2004

PATENT APPLICATION: US/10/762,769

TIME: 16:31:11

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\09202004\J762769.raw

```

212 caacaaagcc gagggcacct actggtctcc aaatcggggg cagcaggagg catggggcgcc 300
213 catggagggg gcttagggga accggtcgat aattggatca agaagctact cgttggtgtc 360
214 gcggcggcgt acatcggttt ggtcgtgctg gtgcccttcc tgaatgtctt cgtccaggcg 420
215 ttcgccaaag gcatcattcc ctctctggag cactgcgcgg acccggaact tctgcacgca 480
216 ctcaagatga cgctgatgct ggcgttcctg acggtgccgc tcaacacggg gtttggcacg 540
217 gtggccgcga tcaacctcac gcgcaacgag ttccccggca aggtgttctt gatgtcgctg 600
218 ctggacctgc ctttctccat ctgcctcggt gtgactggcc tgatgtctac gctgctgtac 660
219 ggccgcaccg gctgggttcgc ggcgctgctg cgggagaccg gcatcaacgt ggtgttcgca 720
220 ttcacggggc tggccctggc caccatgttt gtgacgctgc cgttcgtggg gcgcgagctg 780
221 atccccatcc tggagaacat ggacctgtcg caggaggagg cggcgagaac gctggggggc 840
222 aacgactggc aggtgttctg gaacgtgacg ctgccaaca tccgctgggg cctgctgtac 900
223 ggcgtgatcc tgtgcaacgc ccgagccatg ggcgagttcg gagcctgtgc cgtcatctcg 960
224 ggcaacatca tcggccgcac gcagacgctg acgtgttctg tcgagtcgc ctacaaggag 1020
225 tacaacacgg aggcggcggt cgccggcggt gtgctgctga gcgcgctggc gctgggcacc 1080
226 ctgtggatca aggacaaggt ggaggaggcg gcggcgcgag agagccgcaa gtagagagga 1140
227 gcaggcgcg cggcagcggt cggcagtggc agcggcagcg gcggagagcg gcagctggag 1200
228 aggagcaggc ggtggcgcg gcgagcgga aatagagagg tgcagcaagg aggcaggcgc 1260
229 cgacgcgagg ggaggcgctg gtggtgggtg tgcgtgggtg cttggtccgt ggccagggtg 1320
230 cctggcctgg gtagttggtg tgtgggtgaa gctgattcct gtttgggtga ggccggcgag 1380
231 ttctgaagg aagcaaggaa ggacagtgcc gcagtacca gcgggtaatg gtaaggagag 1440
232 tgacacgtgt ggcgttctgt tgcgtgctgc cgcagtctta acgcagcggg agcagcttct 1500
233 ctgtctgatg tctaaccggg gcgttgatg ctgataatag acggagggcg aaggagagag 1560
234 gcgcggttca gatggggtaa aagctgttgg aaatcaacac gtgcagcggg tgggttgcat 1620
235 ttgtgatcac tggacgttct gagtgttccg tgcgcctata gcgcgtgctg tgcataata 1680
236 cgccgcgcgg cgcataaaac atgactgcat gtgtcggtgt tgacgggtaca gttatgccgt 1740
237 gccccgtttt acaagcggga tagaggcaca ctccacgtag tatgcattga gccagtaga 1800
238 ctctggtcag aaggccggta aatttacatg tgcgtggtg aacctgtaa gtcattggcc 1860
239 aag
241 <210> SEQ ID NO: 5
242 <211> LENGTH: 2253
243 <212> TYPE: DNA
244 <213> ORGANISM: Chlamydomonas reinhardtii
246 <400> SEQUENCE: 5
247 gtacttcaat tgtcagaatg gcgtcgctgc tcgtcaaac aacatcgcg cttggcgctc 60
248 gccagctgc gcaagctggc cctgtcgccc aaatggcacc gatggcaagc cgagtgcagc 120
249 cggcgatgcc tagcgcgctg ctccactgc acgccagagc gacaacaact tcagtgcctt 180
250 gccgggcagc cagcatcgac aaacctgtcg ttactactcc tcgagattcg tcgcaacagt 240
251 cctccaatgg ggcaggagaa gtgtccatgt ccatacatc catggacgag gttggaccct 300
252 cttatgaggg aatcattaca gacgcgccta cacgaccaac ggggctttat gtgcgggtgc 360
253 gcaacatggt gaagcacttc agcaccgcca aaggcctgtt cagggcggtg gacggcgtgg 420
254 acgtggacat cgagcccagc tccatcgtag cgctgctggg gccagcggc agcggcaaga 480
255 ccacattgct gcgcctcatt gcaggcctgg agcagcccac gggcggaac atctactttg 540
256 acgacacgga cgcgaccaac ctgtccgtcc aggaccgcca gatcggttc gtgttcaga 600
257 gctatgcgct gttcaaccac aagacagttg cggagaacat caagtttgga ctggaggtgc 660
258 gcaagctcaa catcgaccac gacaagcgcg tggcgagct gctggcgctg gtgcagctca 720
259 ccggcctggg ccaccgctac ccgcgccaac tgcgggcggt gcagcggcag cgtgtggcgc 780
260 tggcgcgcg cctggcctcc aaccgcggc gtgctgctgt ggacgagccc tttggcgcg 840
261 tggacgcggt ggtgcgcaag cagctgcgca gggggctgcg cgagatcgtg cgcagcgtgg 900
262 gcgtgaccac catcattgtg acgcacgacc aggaggaggc gttcgacctg gcggacaagg 960

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 09/20/2004
PATENT APPLICATION: US/10/762,769 TIME: 16:31:12

Input Set : D:\Seqlist.txt
Output Set: N:\CRF4\09202004\J762769.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; Xaa Pos. 438

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/762,769

DATE: 09/20/2004

TIME: 16:31:12

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\09202004\J762769.raw

L:445 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:432